REMARKS

Claims 1-12 and 14-25 are pending. Claims 1 and 14 are amended herein.

The Examiner has rejected the claims under 35 U.S.C. § 103(a) according to the following table:

Claims	References
1-2, 4-9, 12, 14-15, 17-22, and 25	Borella, Harris, and Scott
3, 16	Borella, Harris, Scott, and Anandakumar
10-11, 23-24	Borella, Harris, Scott, and Orleth

Although Applicant disagrees with each of these rejections, Applicant has amended the claims to make it clear that the receiving endpoint performs each of the steps of the recited method, including detecting when either the buffer has reached a predetermined threshold or a burst has ended and playing back the audio data. Applicant respectfully traverses these rejections below.

Borella describes a system for sending voice data similar to telephone conversations over a non-guaranteed network medium such as the Internet. The system in Borella sends multiple copies of the data from the sender to the receiver, with each copy having differing characteristics such as sampling rate and whether error-correction is present. The receiver places these multiple copies in separate buffers and then periodically examines the data that it has received to see which buffer is the highest quality buffer that is complete enough to be played from. The Examiner relies on Borella for teaching buffering audio and detecting talk spurts.

Harris is primarily concerned with the power level that is required to transmit data reliably in a wireless network, such as a cell phone network. The Examiner relies on Harris for teaching playing audio when a predetermined threshold has been reached and playing audio when a burst has ended. However, the environment described in Harris relies on a cell phone carrier, called the infrastructure throughout the specification, to make decisions as to when to play audio and when to stop playing audio and store it in a buffer, rather than

the receiving endpoint (i.e., the cell phone). Harris states "when infrastructure 124 wishes to instruct, or authorize, MS 104 [(i.e., mobile station or cell phone)] to begin playing out audio information, the infrastructure inserts a value of '1' into the AUDIO_CTRL data field of an RLP frame." Harris, 10:47-50. RLP is a protocol used to transmit information from the infrastructure to the cell phone, and the AUDIO_CTRL data field is part of the protocol that instructs the cell phone to play audio data when the value is one and to buffer audio data when the value is zero.

Scott describes a system for managing jitter in which the receiver receives both voice and silence packets from the sender. The system in Scott attempts to maintain a target jitter buffer size by either inserting a new silence packet if the jitter buffer is too small, or discarding a received silence packet if the buffer is too large. Figures 10 and 13 of Scott demonstrate this system. While Figure 13 of Scott labels the beginning of two bursts, Scott does not disclose any method which either detects the end of a burst or uses such information as a cue to begin playing back audio data in the buffer. Rather, Scott is always playing back audio data from the jitter buffer, "[t]he advantages of the present invention are provided by the ability of the jitter buffer manager 320 to maintain jitter buffer 330 in such a way that the outputted traffic is <u>continuous.</u>" Scott, 7:66-8:2. The Examiner relies on Scott for teaching determining accumulated jitter in a previous burst and waiting for a silent period based on the accumulated jitter.

In contrast, Applicant's technology is directed to reducing the amount of jitter in bursty audio. When data is transmitted over a network such as the Internet, the packets take varying amounts of time to arrive. If audio is played back as it is received, or is buffered and then played back at the rate at which it was received, it will sound choppy to a human listener due to jitter. Applicant's technology reduces this effect by removing the jitter from audio as it arrives, and then making up for the time consumed by the jitter by increasing the periods of silence in between talk bursts. On the other hand, if audio is buffered for too long, then speech will have a noticeable delay that makes a two-way conversation difficult. Therefore, Applicant's technology plays audio from the jitter buffer

both when a burst has ended (to make sure enough buffering occurs) and when a predetermined threshold has been reached (to avoid buffering for too long). Claim 1 recites "upon detecting that the buffer contains an amount of audio data which matches a predetermined threshold amount, playing at the receiving endpoint the audio data contained in the buffer" and "upon detecting that a burst has ended, at the receiving endpoint: playing the audio data contained in the buffer." Claim 14 recites "upon detecting that the buffer contains an amount of audio data which matches a predetermined threshold amount, playing at the receiving endpoint the audio data contained in the buffer" and "upon detecting that a burst has ended, playing at the receiving endpoint the audio data contained in the buffer." Thus, each of Applicant's claims recites playing audio when a predetermined threshold is detected at a receiving endpoint and playing audio when the end of a burst is detected at a receiving endpoint.

The combination of Borella, Harris, and Scott do not teach playing audio when a predetermined threshold is detected at a receiving endpoint and playing audio when the end of a burst is detected at a receiving endpoint. While Borella detects talk bursts, it uses this information only as an indicator of a good time to switch among its various buffers. Borella does not change the duration of the periods of silence in between talk bursts, but rather is always playing from one of its buffers. Moreover, the method described by Harris (which makes decisions on when to play audio at infrastructure that is an intermediary to the sender and receiver) differs substantially from Applicant's technique which makes decisions on when to play audio at the receiving endpoint. Such decisions are made at the receiving endpoint because additional jitter will be introduced into the stream of packets flowing from the infrastructure to the mobile station in Harris, and therefore Harris will suffer from the same limitations noted by Applicant in the present application's background which describes the jitter introduced when transmitting audio data, such as speech, from one endpoint to another on a network. By managing the receiver's buffer at the infrastructure rather than at the receiving cell phone, Harris is unable to eliminate the jitter introduced during the transmission from the infrastructure to the receiving cell phone.

Thus, Applicant's claims recite a novel and nonobvious combination of elements that is neither taught nor suggested by the combination of Borella, Harris, and Scott. Accordingly, Applicant respectfully requests that these rejections be withdrawn.

In view of the above amendments and remarks, Applicant respectfully requests reconsideration of the present application and its early allowance. Applicant believes no fee is due with this response. However, if a fee is due, please charge our Deposit Account No. 50-0665, under Order No. 418268890US from which the undersigned is authorized to draw.

Dated: March 8, 2006

Respectfully submitted,

By Maurice J Pirio

Maurice J. Pirio
Registration No.: 33,273
PERKINS COIE LLP
P.O. Box 1247
Seattle, 98111-1247
(206) 359-8000
(206) 359-7198 (Fax)
Attorney for Applicant